Non-stationary flood frequency analysis in continental Spanish rivers using climate and reservoir indexes as external covariates

Jesús López (jeslpede@posgrado.upv.es), Félix Francés (ffrances@hma.upv.es)

Research Group of Hydrological and Environmental Modelling (http://lluvia.dihma.upv.es/)

Research Institute of Water and Environmental Engineering (IAMMA), Universitat Politècnica de València (Spain), Camino de Vera s/n, 46022, Valencia, España

INTRODUCTION

Traditionally, **stationary** has been the cornerstone of the **"flood frequency analysis" (FFA)**. Recently, the hypothesis of stationarity in the hydrological variables is being questioned (e.g. Milly et al., 2008).

Two factors impact in the violations of the stationarity in the flood regime:

❑ **Climatic** (climate variability)
❑ **Anthropogenic** (changes in land use/land cover and river regulation by dams)

Problems in the classical FFA: Not represent the dynamic of nature.

| Stationary world? Incorrect assumption | → | Non-stationary world? It is necessary and appropriate |

How to address this problem?
Use **new statistical models and procedures** that capture the evolution of pdf's over time.

Which external forcing can be used as covariates in the modelling annual floods?
❑ Climate indexes (large scale atmospheric circulation patterns)
❑ Reservoir index

Objective: Develop a framework of **FFA** under non-stationary conditions based on the GAMLSS models and use climate and reservoir indexes as external drivers in a covariate analysis to describe the variability of annual floods in Spanish rivers.

GAMLSS

GAMLSS models were proposed by Rigby and Stasinopoulos [2005] and provide a flexible modelling framework for annual flood peaks. In a GAMLSS models the observations y_i (for i=1, ..., n) are assumed to be independent with a distribution function $F_Y(y_i | \theta_i)$, where $\theta_i=(\theta_{i1}, ..., \theta_{mi})$ represents a vector of m distribution parameters accounting for location, scale, and shape.

GAMLSS involve several important models, relating the distribution parameters with explanatory variables through a semi-parametric additive model (Eq.2), where if we don't have covariates, the models are reduced to linear parametric models (Eq.3). On the other hand, if we have covariates in the models of Eq. (2), these can be extended including nonlinear parametric terms of the distribution parameters (Eq.4). If $J_k=0$ in Eq.(4) we have nonlinear parametric (Eq.5).

$$g_k(\theta_k) = \eta_k = X_k\beta_k + \sum_{j=1}^{J_k} h_{jk}(x_{jk}) \quad (2)$$

$$g_k(\theta_k) = \eta_k = X_k\beta_k \quad (3)$$

$$g_k(\theta_k) = \eta_k = h_k(X_k\beta_k) + \sum_{j=1}^{J_k} h_{jk}(x_{jk}) \quad (4)$$

$$g_k(\theta_k) = \eta_k = h_k(X_k\beta_k) \quad (5)$$

Probability Density Functions (PDF`s)

We selected 5 distributions usually used in floods frequency analysis: *Gumbel (GU), Lognormal 2p (LNO), Weibull (WEI), Gamma 2p (GA) y Generalized Gamma (GG)*.

Model Fitting Criteria

To avoid model over-fitting and in agreement with the parsimony principle, stepwise methods were used with respect to both **Akaike Information Criterion** (AIC) and **Schwarz Bayesian Criterion** (SBC), equations 6 and 7 respectively. The use of these criteria represents a tradeoff between model complexity and precision.

$$AIC = -2\ln(ML) + 2k \quad (6)$$

$$SBC = -2\ln(ML) + k\ln(n) \quad (7)$$

The goodness of fit of the selected models was verified in accordance with the recommendations of Rigby and Stasinopoulos [2005], checking the normality and independence of the residuals.

DATA & MODELS

Case – study: Continental Spain

Basins:
❑ 8 natural
❑ 12 altered

Low Altered regime

High Altered regime

Reservoir index:

We evaluate the hydrologic effect of dams by means of a dimensionless reservoir index (RI) defined as:

$$RI = \sum_{i=1}^{N} \left(\frac{A_i}{A_T}\right) \cdot \left(\frac{C_i}{C_T}\right) \quad (1)$$

N number of upstream dams
A_i drainage area of each dam
A_T is the drainage area of flow gauge station
C_i reservoir capacity
C_T mean annual runoff at the flow gauge station

Reservoir index values

6 flow gauge station exhibit a higher dam impact ($RI > 0.25$).

Flow Gauge Station	RI_{MAX}	Flow Gauge Station	RI_{MAX}
2002	0.276	5047	0.991
2015	0.014	7006	0.360
2052	0.058	7029	0.465
4014	0.517	8032	0.013
5004	0.329	9002	0.009
5029	0.019	9096	0.129

Climate indexes:
❑ North Atlantic Oscillation
❑ Arctic Oscillation
❑ Mediterranean Oscillation
❑ Western Mediterranean Oscillation

Monthly indexes (AOi, MOi, NAOi and WeMOi)
Monthly climate index corresponding to month annual flood occurrence and considering 2 lag months

Winter indexes (AOw, MOw, NAOw and WeMOw)
Average winter climate indexes from the period December to February

Principal Components
CP1 (67% variance)
CP2 (26% variance)

Models Covariates:

First approach: using only the time (t) as covariate

| Modelling annual floods without external forcing |

Second approach: Using climate and reservoir indexes as covariates

| Modelling annual floods with external drivers |

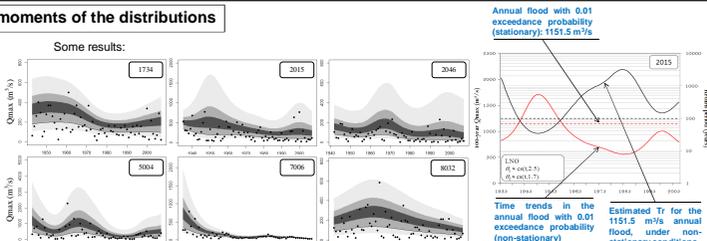| Model 1 | Monthly climate indexes and reservoir indexes |
| Model 2 | Winter climate indexes and reservoir indexes |
| Model 3 | Principal components taken from the EFO's analysis of the winter climate indexes and reservoir indexes |

RESULTS

Modelling annual floods with time-trends in the moments of the distributions

The results are shown by six different quantiles (**0.25, 0.50, 0.75, 0.90, 0.95 and 0.99**).

The plots shown can be considered as a general representation of all results, we show 2 cases in **natural regime (1734 y 2046)**, 2 in **low altered regime (2015 y 8032)** and 2 in **high altered regime (5004 y 7006)**.

It is clearly the presence of **non-stationarities** in the interannual flood regime in Spanish rivers.

Decreasing trends are observed in most flood records in natural regime and low altered since the early 1960`s to late 1990`s, as well as an increasing trends is also evident in the period after 1990's.

Annual flood with 0.01 exceedance probability (stationary: 1151.5 m³/s)

Time trends in the annual flood with 0.01 exceedance probability (non-stationary)

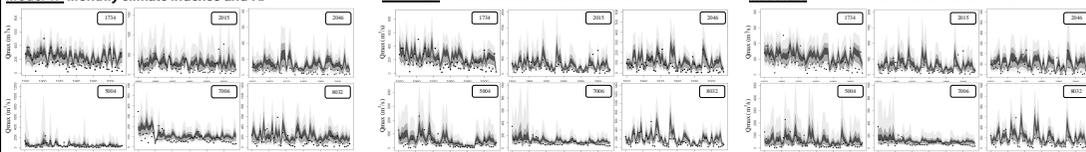Estimated Tr for the 1151.5 m³/s annual flood, under non-stationary conditions

Modelling annual floods with external drivers: climate and reservoir indexes as covariate

Modelling results in six flow gauge station for the 3 models addressed. The results exhibit the regions between the 0.25 and 0.99 quantiles (0.25, 0.50, 0.75, 0.90, 0.95 y 0.99)

Model 1: Monthly climate indexes and RI

Model 2: Winter climate indexes and RI

Model 3: PC`s of winter climate indexes and RI

Stationary and non-stationary models: The results are based on the fitted GAMLLS models for the 2015 flow gauge station.

Model 1 vs Stationary model
❑ Most complex models.
❑ Climate indexes are significant covariates in θ_1 and θ_2.
❑ RI is significant covariate in 6 models (particularly in southern basins).
❑ Dependence between distribution parameters and external covariates are modeled by mean of non-parametrics formulations.
❑ The results seem to adequately reproduce the behavior exhibited by the floods.

Model 2 vs Stationary model
❑ Climate indexes are significant covariates in θ_1 and θ_2 exhibits be independent of external forcings.
❑ The results for Model 2 are less noisy than for the Model 1, and seem to better reproduce the behavior exhibited by the floods.
❑ In 1970-1990 the annual floods experiment a lesser frequency and magnitude.
❑ In 1955-1970 and 1995-2005 the annual floods experiment a higher frequency and magnitude.

Model 3 vs Stationary model
❑ Using the PC`s from the FOES analysis we obtain the most parsimonious models.
❑ The models reproduce better the non-stationarities in the annual flood regime.
❑ Gamma 2p, Generalized Gamma and Lognormal 2p result the best distribution to fit the annual flood data.
❑ The models exhibit significant differences with respect to stationary model.
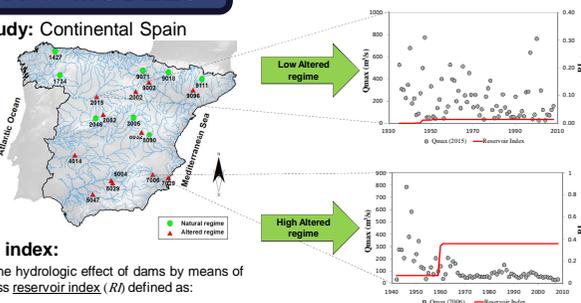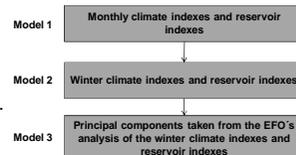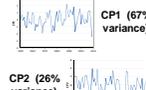
GAMLSS models as predictive tool?
Modeling results of annual floods for the 2015 flow gauge station in the period 1935-1990 using the Model 3. The modeling results are represented by solid lines. The model is then used as a predictive tool (dotted line) and the observation not included in the fitting of the model are represented by gray circles.

Annual floods with 0.01 annual exceedance probability under non-stationary conditions (red line) and stationary conditions (dotted red line). The estimated return period with 0.01 annual exceedance probability (black line) derived from the period 1950-2007 in the 2015 flow gauge station.

CONCLUSIONS

1. The presence of **increasing and decreasing trends** in the temporal evolution of the flood regime, **can be associated to changes in the climate indexes**, particularly in the observed in the winter indexes and the CP`s.
2. GAMLSS provide a useful modelling framework for representing **non-stationarities** exhibited in annual flood records.
3. Results show climate indices **can be incorporate as external covariates** in the non-stationary models.
4. As expected, **AO and NAO** tend to be a **significant predictors** for the majority of the cases in continental Spanish rivers.
5. RI results seem to adequately **reproduce sudden changes caused by dam construction**.
6. A proper estimation based on the non-stationary flood frequency model shows that any flood quantile **varied dramatically during the flood record**.

Acknowledgements

References

- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., et al., (2008). "Stationarity Is Dead : Whither Water Management ?" Science, 573-574.
- Rigby, R., and Stasinopoulos, D. (2005). "Generalized additive models for location, scale and shape." Journal of the Royal Statistical Society: Series C (Applied Statistics), Blackwell Publishing Ltd, 54(3), 507-554.
- Stasinopoulos, D., and Rigby, R. (2007). "Generalized additive models for location scale and shape (GAMLSS) in R." Journal of Statistical Software.